GEOPHYSICAL SIGNAL RECOGNITION

C. H. Chen

Electrical Engineering Department
Southeastern Massachusetts University
North Dartmouth, Massachusetts 02747

ABSTRACT

Progress of pattern recognition as applied to
various geophysical signals is reviewed. They
include the teleseismic signals for nuclear
detection and monitoring, exploration seismic
signals for oil prospecting, marine seismic
signals for study of subsurface structure of
ocean floor, seismic signals for earthquake pre-
diction, seismic signals for intrusion-detection,
and geomagnetic signals for study of irregular-
ities of the earth's magnetic fields. With the
exception of teleseismic signal, research in
geophysical signal recognition is still in its
infancy. We expect to see considerable research
progress in the next decade.

1. INTRODUCTION

In recent years, computers have played an in-
creasingly important role in geophysics especial-
ly in seismic studies such as the petroleum
exploration, mineral extraction, nuclear detec-
tion, earthquake research, marine seismic pro-
filing, and intrusion detection, etc. Computers
are needed to process large volumes of seismic
data from which useful information must be ex-
tracted accurately. In addition to the seismic
data, geomagnetic signals are also processed by
computer in large quantities. A primary function
of computers has been in filtering and storage of
the geophysical data while the recognition and
interpretation of the filtered results is performed
mainly by geophysicists or human operators. How-
ever efforts have been made by many researchers to
use pattern recognition techniques in geophysical
signal processing and recognition with some promis-
ing results. In this paper we shall examine the
progress in geophysical signal recognition and the
future directions of such research. An extensive
list of references provided as bibliography at the
end of the paper clearly indicates the amount of
research activities in this area in the past ten
years.

Pattern recognition techniques considered in this
paper include preprocessing, feature or primitive
selection, syntactive analysis, decision making
and learning processes. Among all geophysical
problems, seismic discrimination clearly can be
formulated as a pattern recognition problem and
has been fairly extensively studies in this context.
For the exploration seismic data from which

detection is made of petroleum, mineral or natural
gas, pattern recognition can be used as computer-
aided interpretation. Similarly for earthquake
study and seismic intrusion detection, pattern
recognition provides an additional dimension to
problem solving. In other geophysical signals,
the need for pattern recognition is less evident.
As different geophysical signals require different
processing and recognition approaches, the recog-
nition problems of various geophysical signals
will be considered individually.

2. RECOGNITION OF TELESEISMIC SIGNALS

Teleseismic signals are the seismic waveforms
resulting from nuclear explosions and earthquakes
which may originate from unknown locations.
Shallow earthquakes are easily confused with
nuclear explosions. Seismic networks throughout
the world coupled with a priori knowledge about
explosion and earthquake events have made it
possible now to detect the nuclear explosion with
good accuracy. Automatic pattern recognition is
useful to aid or verify the human interpretation
at the present stage of development. Currently
available recognition results have not been close
to error-free to make the automatic recognition
reliable.

Mathematical feature extraction by using dynamic
spectral ratio, complexity, multivariate auto-
regressive parameters in conjunction with the
Karhunen-Loeve expansions have been considered for
seismic discrimination (e.g. [6][9][19][20][31]).
Features with real geophysical meaning appear to
be quite important for seismic discrimination.

Examples are [19]:

1. Earthquakes produce approximately equal amounts
of P and S waves, while explosions produce more
P waves.

2. Earthquakes have relatively deep foci, while
explosions have only shallow foci.

3. Earthquakes give anaseismic and kataseismic
first onsets while explosions give anaseismic
first onset everywhere.

4. The duration of wave trains is shorter for
explosions than for earthquakes.

Unfortunately the direct application of these
discriminating features is not possible. Although

81 10 21

explosions provide more P waves than S waves, not all earthquakes follow the average pattern. Further it is very difficult to identify particular phases for lower magnitude earthquakes. Focal depths hold a great deal of promise in discrimination. However there is still too much error associated with measurement to warrant its exclusive use. The difficulty in determining the direction of onset, due to noise, obscures the anaseismic-kataseismic discrimination feature. Again, the signal-to-noise ratio is too small, for small-amplitude events, to discriminate using the duration of the wave trains. An interesting discussion of geophysical features for nuclear detection is given by Bolt [3]. For automatic classification, the features must be computable to a good accuracy. This requires both digital signal processing and statistical techniques.

Figure 1 shows some typical seismic records which we have studied along with their Fourier amplitude spectra. The sampling rate is 10 samples/sec. with 1200 samples per record. Based on our extensive comparative study of various feature selection criteria the spectral features when accurately computed can perform the best with simple classifier. For example, the high-resolution maximum entropy spectral analysis can provide more effective features than those available from the conventional modified periodogram method using fast Fourier transform. Figure 2 shows the maximum entropy spectra of two explosion and two earthquake events for a duration of 200 samples from sample number 701 to 900. For the explosion the spectral ratio defined as the ratio of signal energy above 0.5 Hz to that below 0.5 Hz tends to be greater than 1 while for the earthquake such ratio tends to be less than 1. The nearest neighbor decision rule is used for classification. For each pattern class, 40 seismic records are available for learning. A total of 243 seismic records which do not include learning records is for classification. Spectral ratio for 200 samples from sample number 550 to 749 and spectral ratio for 200 samples from sample number 750 to 949 are employed as two features in classification. The final classification result is 14 errors out of 243 for 94.24% correct recobnition. This is the best available result with two features.

Another important feature extraction technique is the short-term spectral features. A 64-point window slides through th 1024-point seismic record generating 64 time segments with 48-point overlap between the adjacent segments. For each 64-point segment, compute the Fourier transform. The first 32 points of the Fourier transform correspond to the 5-Hz frequency interval of the seismic wave. Define P_1 = signal power covering the first to fifth spectral points (0-0.78 Hz band), P_2 = signal power covering the sixth to 14th spectral points (0.78-2.18 Hz band), and P_3 = signal power covering the 15th-32nd spectral points (2.18-5.0 Hz band). Each P_i, $i = 1, 2, 3$ can be plotted as a function of time, i.e., the segment number. For each curve the ratio of the signal energies between the first 32 segments and the second 32 segments is called a short-time spectral feature. For each seismic record, there will thus be three feature values

which form a vector. A set of 50 learning (training) records is taken from each class. Using the well known Euclidean distance nearest-neighbor decision rule the best performance is around 80% correct recognition on test set. For each class, a covariance matrix V_i, $i = 1, 2$ can be determined from the training set. A modified procedure is to compute the squared distance

$$(x - P^{(i)}) \, V_i^{-1} (x - P^{(i)}) = d_i^2 \, , \; i = 1, 2 \quad (1)$$

and choose the class which provides the minimum d_i^2. Here $P^{(i)}$ is a nearest neighbor belonging to the ith class and x is the three-dimensional feature vector representing a seismic record. By using such weighted Euclidean distance nearest neighbor decision rule, a 94.17% correct recognition (13 errors out of 223) is obtained. If the Bayes decision rule is used under the assumption of multivariate Gaussian density for each class, the performance is 87.92% correct recognition (27 errors out of 223). This experimental result indicates that the stationary probability assumption of the seismic record is not good. Other work on autoregressive feature extraction also based on the assumption of stationary random process for the seismic record has provided less recognition performance. The short-term spectral features described above have taken the nonstationarity into consideration.

Among other computer studies, the features based on the complex cepstral point followed by Fisher's linear discriminant gave a 94.4% correct recognition results for 36 seismic records. The autocovariance function also can provide around 89% correct recognition. The results are somewhat depending on the pre-filtering operations used. For a single discriminant, the ratio of surface wave magnitude M_s to the bodywave magnitude M_b is most effective for events with $M_b > 4$ as the shallow earthquakes typically produce more surface waves than do explosions of equivalent energy. The combined use of statistical and structural features [7] in conjunction with an optimized decision tree is a very promising approach. Syntactic method of teleseismic signal recognition has also been examined. As a highly reliable and fully automatic seismic recognition system is not likely to be available in the near future, interactive pattern recognition is more practically feasible and some important effort has been made in this direction [21].

3. EXPLORATION AND MARINE SEISMIC SIGNALS

In the exploration seismic study detection is made of petroleum or natural gas. In the marine seismic study, an accurate mapping is made of subsurface geological structure of the ocean floor. Although the purposes are quite different, the two seismic signal studies are similar. As the new petroleum deposits are much more difficult to discover on land now, exploration of petroleum deposit in the ocean floor is increasingly important and the demand for marine seismic study for oil exploration is very high. For both seismic signals, the major efforts made in the past have been in

digital filtering and estimation of subsurface parameters. Pattern recognition problems can be formulated for analysis and detection of the seismic signals. In marine seismic profiling [5], for example, the physical parameters which describe the structure of the deep sea floors correspond to feature sets in pattern recognition. Examples of pattern classes are various rocks, sediments, mantles and oceanic layers of low, medium and high porosity. In the exploration seismic study feature extraction and decision theoretic approaches have been examined [15][16]. An important recent development is the syntactic approach [26] to segment the one-dimensional logging data, i.e., measurements of various physical rock parameters from transducers inside the borehole. The transition patterns between signal blocks are described by a set of grammatical rules.

As compared with the teleseismic signals, the exploration and marine seismic signals are much more stationary and thus both statistical and syntactic pattern analysis can be quite effective for properly formulated recognition problems.

4. MINERAL EXTRACTION AND EARTHQUAKE PREDICTION

The use of seismics in mining exploration has been restricted so far to only some trial measurements. If seismic methods are to become popular in ore exploration, we can be sure that the signals of interest will be hidden far more "deeply" than is the case for the signals in oil propsecting. The highly developed reflection seismology that reveals subsurface structure information will be useful to mineral extraction. Pattern recognition techniques useful for petroleum exploration should also be suitable for mineral extraction.

Computer-aided techniques have been quite helpful in earthquake prediction. However pattern recognition techniques have not been employed in earthquake study, to our knowledge; but the potential is there. For example, automatic processing algorithms for microearthquake data have been examined [1], based on the syntactic structural analysis of waveforms [27][28]. There is always ambiguity associated with measuring the first arrival time from seismograms whether it is done by a seismologist or by a machine since the seismic signals are of unknown shape and are contaminated by noise. Such ambiguity can be reduced by combining the processes of picking arrivals and locating the events in an iterative fashion. A seismogram can be described as "noise followed by seismic signal", where both "noise" and "seismic signal" would have grammatical rules describing their structure. Current picking algorithms [2] use this type of structural information but in a rather ad hoc manner.

In summary, application of pattern recognition to earthquake prediction is in its infancy. Obviously various structural and geophysical information must be combined to derive a feasible algorithm for automatic prediction and interpretation.

5. INTRUSION-DETECTION USING SEISMIC SENSORS

One of the automatic intrusion-detection techniques is to use the seismic sensors that detect the "footstep signal" in the presence of usually strong correlated noises due to vehicle motions in the immediate vicinity. In this case the real signal is impulse-like with unknown arrival time and duration. It is necessary to employ noise cancelling algorithms such as the adaptive digital filtering and the Kalman filtering to suppress the background noise so that the footstep signal can be enhanced [10][11]. Figures 3 & 4 show two typical sections of seismic data containing footstep signals, along with the results of adaptive digital filtering and adaptive Kalman filtering. A major portion of background noise has been removed after filtering even if the noise is very strong. Each section as shown has 960 samples at a sampling rate of 200 samples/sec. For this kind of seismic signal, spectral features are not suitable for detection. The recognition or detection problem here is to determine for each section of seismic data whether the footstep signal is present or not. The correlation detection or matched filter detection is not useful here. A simple peak detection with properly selected threshold will, however, serve the purpose. There does not appear to have any "structure" in this particular seismic signal.

6. OTHER GEOPHYSICAL SIGNALS

Geomagnetic signals provide some information on geophysical phenomena, which is not available from the seismic data. The spectral analysis is the main computer study that has been performed with the geomagnetic signals. An example of geomagnetic signal is the micropulsations which are often associated with structural disturbance (or storms) in the magnetosphere. Detecting a particular micropulsation event can provide a diagnosis of the properties of the magnetosphere. As the magnetosphere changes because of its varying interaction with the solar wind, the properties of the micropulsation signatures should response. The recognition problem here is somewhat similar to the detection of machine malfunction. As a large volume of geomagnetic signals is received daily, automatic processing and recognition will certainly be an area of future research study.

Other geophysical signals include sunspot numbers, etc. As several different sensors are frequently employed in examining a particular geophysical phenomenon, how to effectively utilize information from all sources is a challenging problem in recognition study.

7. CONCLUSION

As compared with recognition of speech and biomedical signals, progress in geophysical signal recognition is much slower. There are many promising areas of research as described in this paper. Suggested future approaches are the following,
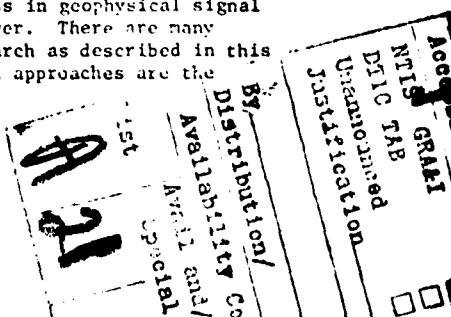
(1) Understanding the geophysical phenomenon is essential as different geophysical signals may require quite different recognition approaches.

(2) Signal processing is an integral part of signal recognition as a number of distinct properties can only be derived after extensive signal processing.

(3) Fully automatic recognition may not be a realistic goal. Some human interaction may be necessary to complete the recognition task.

(4) Information from various geophysical sensors should be integrated and utilized to arrive at the best recognition result.

REFERENCES

1. K.R. Anderson, "Automatic Processing of Local Earthquake Data", Ph.D. Thesis, MIT Dept of of Earth and Planetary Sciences, August 1978.

2. K.R. Anderson, "Automatic analysis of micro-earthquake network data", Geoexploration, Vol. 16, pp. 159-175, 1978.

3. B.A. Bolt, "Nuclear Explosions and Earth-quakes, The Parted Veil", W.H. Freeman and Co., San Francisco, 1976.

4. C.H. Chen, editor, "Computer-Aided Seismic Analysis and Discrimination", Elsevier Scientific Publishing Co., 1978.

5. C.H. Chen, "Application of pattern recognition technqiues to oceanographic signal processing", Record of IEEE International Conference on Engineering in the Ocean Environment, Panama City, Fl. Sept. 1970.

6. C.H. Chen, "Seismic pattern recognition", Geoexploration, Vol. 16, pp. 133-146, April 1978.

7. C.H. Chen, "On statistical and structural feature extraction", in "Pattern Recognition and Artificial Intelligence", ed. C.H. Chen; Academic Press, 1976.

8. C.H. Chen, "On digital signal modelling and classification with the teleseismic data", IEEE ASSP Conference Record, April 1978.

9. C.H. Chen, "Application of pattern recognition to seismic wave interpretation", in Chapter 5 of "Applications of Pattern Recognition", ed. K.S. Fu; CRC Press, Inc., 1981.

10. C.H. Chen, and J. Chen, "Note on Kalman filter-ing for seismic signal detection", Technical Report SMU-EE-TR-80-1, February 1980.

11. C.H. Chen, "Adaptive and learning algorithms for seismic intrusion-detection sensor data", to appear in IEEE Trans. on Pattern Analysis and Machine Intelligence.

12. C.H. Chen, "Digital Waveform Processing and Recognition", CRC Press, Inc., 1981.

13. C.H. Chen, "Digital Processing of Marine Seismic Data", Records of OCEAN'71.

14. O. Dahlwan and H. Israelson, "Monitoring Underground Nuclear Explosions", Elsevier Scientific Publishing Co., 1977.

15. R.J.P. deFigueiredo, Special Issue being prepared.

16. R.J.P. deFigueiredo, Private communication.

17. K.S. Fu, "Syntactic pattern recognition and its applications to signal processing, Proc. of the NATO Advanced Study Institute on Pattern Recognition and Signal Processing, Series E, Applied Science No. 29, Sijthoff & Noordhoff International Publishers, The Netherlands, 1978.

18. K.S. Fu, "Syntactic Pattern Recognition, Applications", Springer-Verlag, 1977.

19. R.B. Ives, "Feature and decision function selection for the detection of nuclear detonations from seismic signatures", Conference Record of the 1976 IEEE Joint Workshop on Pattern Recognition and Artificial Intelligence.

20. R.B. Ives, "Dynamic spectral ratios as features in seismological pattern recognition", Conference on Computer Graphics, Pattern Recognition and Data Structure, 1975.

21. R.B. Ives, presentation given at the First International Symposium on Computer-Aided Seismic Analysis and Discrimination, 1977.

22. J. Lin, unpublished report, Advanced Auto-mation Research Lab, Purdue University, 1977.

23. H.H. Lin and K.S. Fu, "An application of syntactic pattern recognition to seismic discrimination", paper for the Second Inter-national Symposium on Computer-Aided Seismic Analysis and Discrimination, 1981.

24. P.F. Mohajeri, "Some alternate approaches to noise compensation in seismic analysis", MIT Lincoln Lab., Technical Note, 1971-14.

25. A.V. Oppenheim, "Applications of Digital Signal Processing", Prentice-Hall, 1978.

26. H.E. Stephanou, "A syntactic approach to well log segmentation", presented at the 1980 Workshop on Digital Signal and Waveform Analysis.

27. T. Pavlidis, "Structural Pattern Recognition", Springer-Verlag, 1977.

28. T. Pavlidis, "Curve fitting", presented at the 1980 Workship on Digital Signal and Waveform Analysis.

29. E.A. Robinson and M.T. Silvia, "Deconvolution of Geophysical Time Series in the Exploration for Oil and Natural Gas", Elsevier Scientific Publishing Co., 1979.

30. E.A. Robinson, Multichannel Time Series Analysis with Digital Computer Programs", Holden-Day, 1967.

31. R.H. Schumway and A.N. Unger, "Linear discriminant functions for stationary time series", Journal of American Statistical Association, Vol. 67, pp. 948-956, 1974.
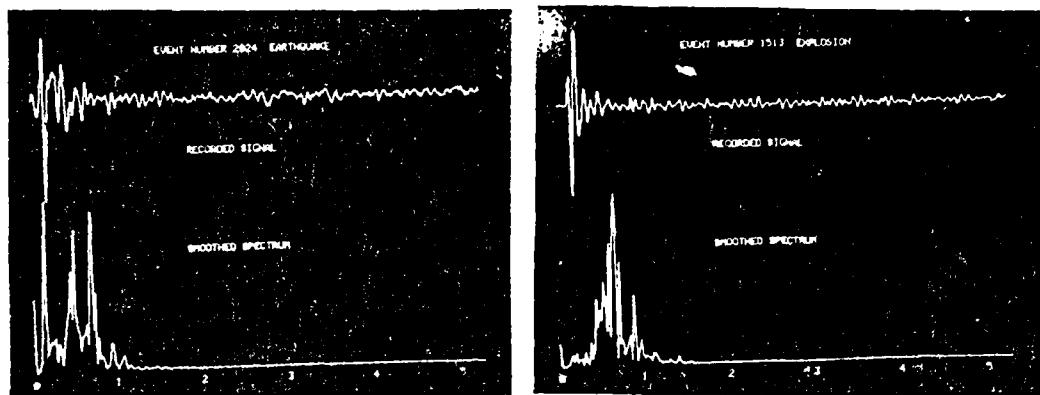
Figure 1
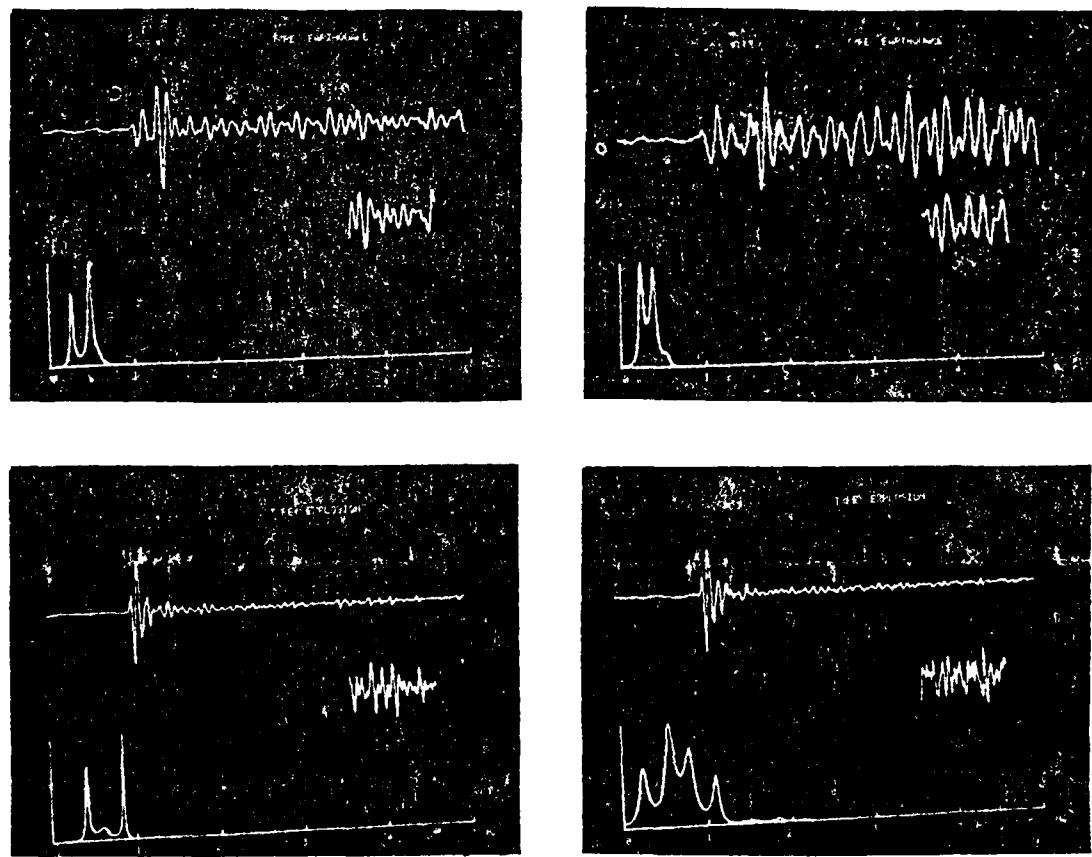
Figure 2
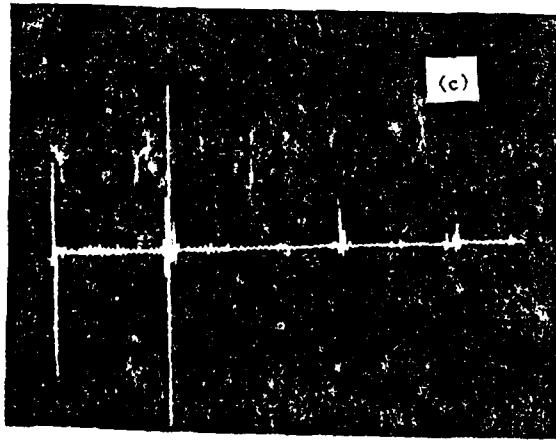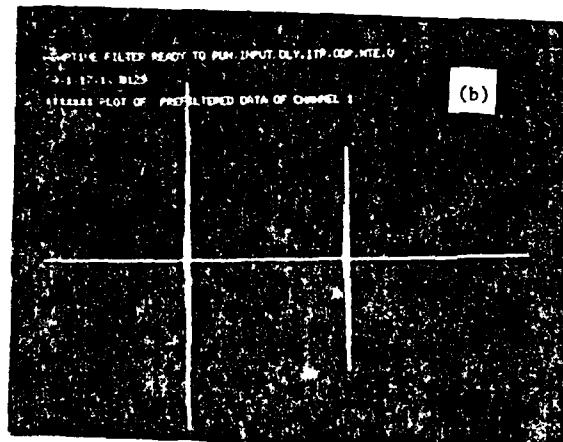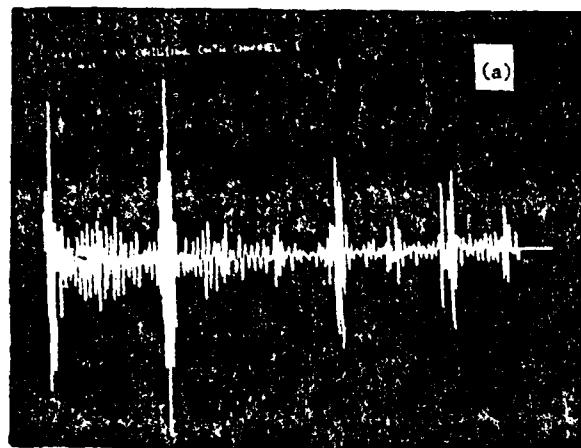
Figure 3 Original wave(a)
        Adaptive digital filtering(b)
        Adaptive Kalman filtering(c)
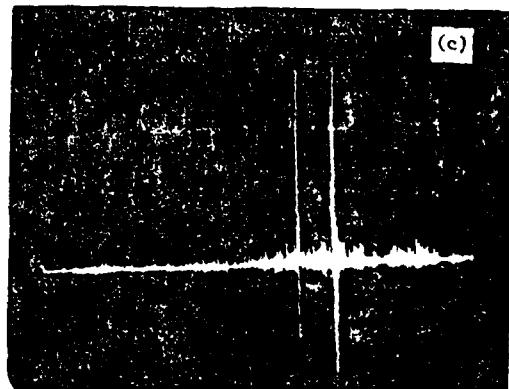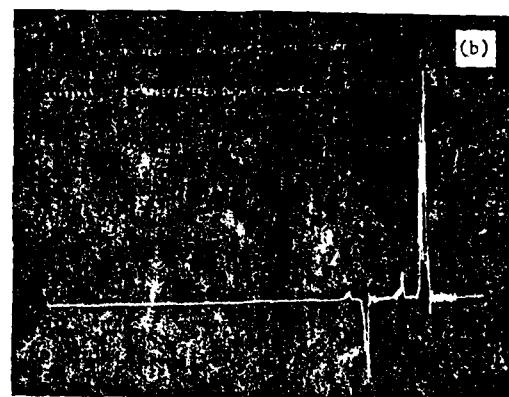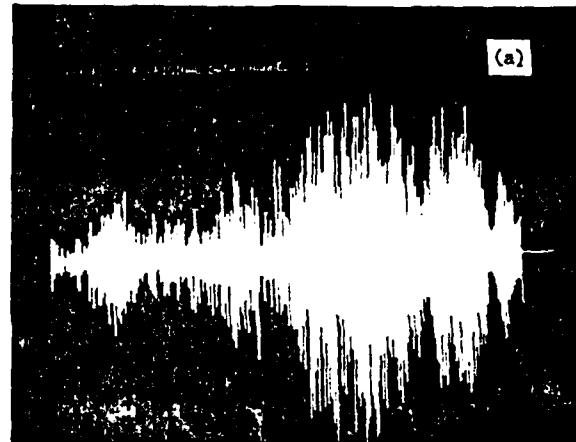
Figure 4 Another original wave(a)
        Adaptive digital filtering(b)
        Adaptive Kalman filtering(c)